

The influence of the assessment method on the incidence of visual field progression in glaucoma: a network meta-analysis

Citation for published version (APA):

Ernest, P. J. G., Viechtbauer, W., Schouten, J. S. A. G., Beckers, H. J. M., Hendrikse, F., Prins, M. H., & Webers, C. A. B. (2012). The influence of the assessment method on the incidence of visual field progression in glaucoma: a network meta-analysis. *Acta Ophthalmologica*, 90(1), 10-19. <https://doi.org/10.1111/j.1755-3768.2010.01995.x>

Document status and date:

Published: 01/02/2012

DOI:

[10.1111/j.1755-3768.2010.01995.x](https://doi.org/10.1111/j.1755-3768.2010.01995.x)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Review Article

The influence of the assessment method on the incidence of visual field progression in glaucoma: a network meta-analysis

Paul J. G. Ernest,^{1,2} Wolfgang Viechtbauer,³
Jan S. A. G. Schouten,² Henny J. M. Beckers,²
Fred Hendrikse,² Martin H. Prins¹ and Carroll A. B. Webers²

¹Department of Epidemiology, Maastricht University, Maastricht, The Netherlands

²University Eye Clinic Maastricht, Maastricht, The Netherlands

³Department of Statistics, Maastricht University, Maastricht, The Netherlands

ABSTRACT.

Purpose: To study and quantify the difference in incidence of progression between methods for the assessment of glaucomatous visual field progression.

Methods: We identified 2450 articles published up to April 2009 in the following data sources: PubMed, EMBASE, and Cochrane. Ten studies covering 30 methods were included. All studies aimed to compare different methods for the assessment of glaucomatous visual field progression in the same study population. A network meta-analysis using a mixed-effects model was performed to combine within-study between-method comparisons with indirect comparisons from other studies. The summarized incidence of progression was calculated for every method, and methods were ranked according to this incidence.

Results: In total, methods were compared in 1040 eyes of 948 patients with glaucoma. On average, 21% of the eyes progressed. When all 30 methods were ranked, the incidence ranged from 2% to 62%. These incidences are corrected for a baseline mean deviation (MD) value of -7 decibels and a mean follow-up time of 6 years. Besides the assessment method, the incidence was only determined by the follow-up period and baseline MD value, leaving no unexplained variance in the incidence of progression.

Conclusion: The incidence of progression varies considerably between different studies. This is mainly caused by the variety of methods used to assess progression but also by differences in follow-up time and baseline visual field loss.

Key words: disease progression – glaucoma – meta-analysis – perimetry

Acta Ophthalmol. 2012; 90: 10–19

© 2010 The Authors

Acta Ophthalmologica © 2010 Acta Ophthalmologica Scandinavica Foundation

doi: 10.1111/j.1755-3768.2010.01995.x

Introduction

Monitoring patients with open-angle glaucoma is based on the reliable

detection of early deterioration of the visual field and subsequent lowering of the intraocular pressure. The aim is to prevent irreversible visual field loss

leading to visual disability (Caprioli 2008). A major problem in the assessment of glaucomatous visual field progression is the lack of a gold standard (Schulzer et al. 1991). As a consequence, investigators often develop a new method for their own purposes. Many different methods have been described, ranging from completely subjective to fully automated calculations (Spry & Johnson 2002; Ernest et al. unpublished data).

If studies are using different methods and subsequently have different results, it is hard to distinguish the contribution of the chosen method from other study characteristics in the quantification of the incidence of progression. This could make studies incomparable to each other. However, the extent to which the use of different methods leads to differences in study outcome between similar studies has not yet been investigated. What has been investigated is the use of different methods in one study population. Most frequently, the mutual agreement of a few methods was investigated by studying the difference in the proportion of patients who develop progression and the overlap in patients classified as having progressed. However, no studies have been published that compare all methods in one study. Therefore, a ranking

of methods is missing, and it remains unclear which methods give a high or a low incidence of progression. The interpretation and the possibility to compare studies will improve with a clear ranking. For example, it would make a real difference to find a high incidence with a low ranked method when compared to finding a low incidence with a high ranked method.

In this study, first, all methods to assess glaucomatous visual field progression used in different studies were compared. Secondly, these methods were ranked according to their incidence of progression, and their difference was quantified using a network meta-analytic approach. Finally, we also investigated the influence of baseline visual field loss and follow-up time and determined the extent to which the incidence of progression was explained by these variables.

Materials and Methods

Search strategy

A systematic computerized search was performed in PubMed, EMBASE, and all databases and registers of The Cochrane Library, in April 2009. The following keywords were used: (glaucoma*) AND (prognos* OR predict* OR progress* OR longitudinal OR cohort OR follow-up) AND (perimetr* OR visual field* OR HFA OR Octopus OR Humphrey (not in author)). The search was limited to articles in English, Dutch, French, or German. A total of 2450 articles were identified.

Selection procedure

All titles and abstracts were screened, and articles were excluded based on predefined exclusion criteria. Of the remaining 782 articles that were studied completely, 48 articles fulfilled the selection criteria (Fig. 1). The selected studies had to follow patients with glaucoma for minimally 1 year with the use of conventional visual field examinations. Studies should also directly compare at least two methods to assess glaucomatous visual field progression within the same study population. After that, the type of perimeter used and a precise description of the methods used for assessment of visual field progression were recorded for these studies. In addition,

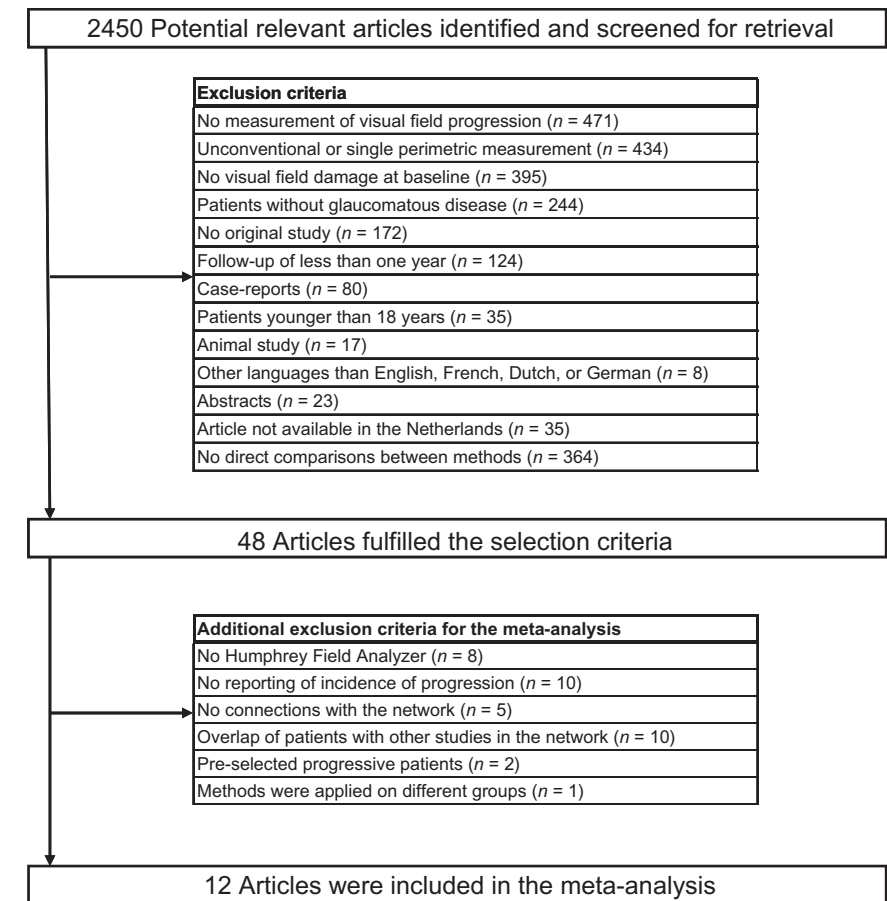


Fig. 1. Flow chart of the article selection process.

the number of participants, duration of follow-up, number of visual field examinations, patient characteristics, incidence of progression, and the publication year were noted.

We combined direct and indirect comparisons of methods in a network meta-analysis. A direct comparison is defined as a comparison of methods in one study using exactly the same study population during the same study period. Indirect comparisons of two methods have not actually been made in one study but are made in different studies by using a common comparator (Bucher et al. 1997). Studies were excluded when no comparisons with other methods in the meta-analysis were possible. This led to the sole inclusion of studies that applied methods based on the Humphrey Field Analyzer (HFA) (Carl Zeiss Meditec, Dublin, CA, USA), because this perimeter was used in most studies on progression and no studies were found that directly compared HFA-based methods with methods from other perimeters. Fur-

thermore, studies that did not report incidence proportions of progression and in which the incidence proportion could not be calculated from the data were excluded. We defined the incidence proportion as the proportion of patients with glaucoma that developed visual field progression during the follow-up period of a specific study. Studies with populations that overlapped with other selected studies were also excluded to obtain a fair comparison between methods. In these cases, the study was chosen that had the most comparisons with methods used in other studies. When overlap of study populations was uncertain, more information was retrieved from the authors. Finally, two studies were found, which investigated a selected group of patients with glaucoma that was considered to be progressive based on clinical judgement (Mayama et al. 2004; Heijl et al. 2008). These studies were excluded, because including them would result in selection bias (Caldwell et al. 2005).

Finally, the data of twelve articles were combined in the network meta-analysis (Fig. 1) (Birch et al. 1995; Smith et al. 1996; Katz et al. 1997, 1999; Katz 2000; Lee et al. 2002; Viswanathan et al. 2003; Artes et al. 2005; Nouri Mahdavi et al. 2007; Arnalich Montiel et al. 2009; Diaz Aleman et al. 2009; Wesselink et al. 2009). Three articles were included that studied different methods in the same population during the same follow-up period (Katz et al. 1997, 1999; Katz 2000). Data of these articles were treated as a single study. The references of all included articles were additionally screened to identify more studies. This did not result in additional articles. However, we found one recent article that described the same study as one of the included articles (Casas Llera et al. 2009). As additional methods were applied in this article, we added these methods to the meta-analysis.

Statistical analysis

If not available, incidence proportions of glaucomatous progression were calculated from the number of eyes that progressed for each method within each study. All incidence proportions were then combined in a mixed-effects model. More specifically, an arm-based network meta-analysis model with dummy variables to identify the various methods was used for the analysis (Salanti et al. 2008). The computations for these analyses were performed with the statistical software R (version 2.10.1). An additional complication in the analysis arose because of the fact that incidence proportions from the same sample using different methods are correlated and therefore cannot be treated as independent observations. The magnitude of the correlation between several incidence proportions from the same sample can only be calculated if a complete cross-classification table of the various methods is provided in an article. Usually, such detailed information was unavailable. We therefore used the method suggested by Riley and co-workers (Riley et al. 2008). This method substitutes a single overall correlation parameter for the study-specific individual correlations.

The model therefore is given by:

$$y_{ij} = \beta_1 M_{1,ij} + \beta_2 M_{2,ij} + \dots + \beta_x M_{x,ij} + u_{ij} + \varepsilon_{ij},$$

where y_{ij} is the observed incidence proportion in study i for method j , $M_{1,ij}$ is a dummy variable (i.e. coded 1 or 0), indicating whether y_{ij} corresponds to method 1 (and similarly for $M_{2,ij}$ through $M_{x,ij}$), the u_{ij} values denote the random effects, and the ε_{ij} values denote the sampling errors. The model assumes that the variances of the u_{ij} 's are equal to τ^2 , denoting the amount of heterogeneity in the true incidence proportions. Furthermore, the correlation between two u_{ij} 's from the same sample (i.e. study) is denoted by ρ . The variances of the ε_{ij} values were estimated based on the observed incidence proportions and can be denoted by v_{ij} . Moreover, two ε_{ij} values from the same study are also assumed to be related to each other with their correlation given by ρ . Finally, assuming normality of the u_{ij} and ε_{ij} values, the parameters of the model were then estimated using restricted maximum likelihood estimation. As a result, we obtained the following estimates from the model: the average incidence proportions for the various methods (estimated values of β_1 through β_x), the amount of heterogeneity in these results (estimate of τ^2), and an estimate of the correlation for incidence proportions arising from the same study (estimate of ρ). The precision of the estimates of β_1 through β_x was quantified via corresponding 95% confidence intervals obtained in the usual manner (i.e. estimate of $\beta \pm 1.96$ SE, where SE denotes the standard error of the estimate of β). A confidence interval was also obtained for τ^2 , based on a profile likelihood method (Hardy & Thompson 1996).

The mean deviation (MD) value at baseline and the mean follow-up time were also included as covariates in the model, so that the incidence proportions from the various methods could be compared while controlling for the influence of these covariates. We included these covariates initially as linear predictors in the model but also examined whether the incidence proportions changed as an accelerating (quadratic) function thereof. Another potential covariate was the mean number of visual field examinations for the patients in each study. How-

ever, this covariate correlated very strongly with the mean follow-up time in a study ($r = 0.79$), and therefore, it was almost impossible to distinguish. We chose to use follow-up time over the number of examinations because its influence on the incidence of progression is inherently more logical.

Where possible, we also compared the incidence proportions of two methods from the network model with incidence proportions of the same two methods, which were based only on direct comparisons. To use a reasonable amount of data for getting more stable results, this additional analysis was only performed for methods that were directly compared in three or more studies.

Results

The flow of articles included in this systematic literature review is shown in Fig. 1. Twelve articles that studied 30 methods in ten studies were included in the meta-analysis. All methods were named and classified in six groups according to their main characteristics (Table 1).

Six groups of methods

The first group consisted of methods based on the glaucoma progression analysis (GPA), which is similar to the visual field endpoint in the Early Manifest Glaucoma Trial (EMGT) (Leske et al. 1999). We abbreviated this group to 'group G', from GPA. The GPA is an event analysis based on pattern deviation values and is included in the new HFA software. When significant deterioration ($p < 0.05$) is seen on the pattern deviation probability maps of the GPA printouts in the same three or more points on three consecutive follow-up tests, the software interprets this as likely progression (Leske et al. 1999). Although some studies performed separate analyses after exclusion of eyes with a baseline visual field "out of range" according to the GPA printout of the HFA, we used the incidences of progression in all study eyes. Group A (A from AGIS) consisted of two methods that were based on the Advanced Glaucoma Intervention Study (AGIS) (The AGIS Investigators 1994). Both the AGIS method and the Collaborative Initial

Table 1. Description of the 30 methods studied.

G	Based on glaucoma progression analysis (GPA)
G1	GPA
G2	GPA using total deviation
G3	GPA with 2 follow-up visual fields
G4	Modified GPA; difference between number of improving points and deteriorating points (≥ 3) repeatable at 3 consecutive examinations
G5	Modified GPA using total deviation; difference between number of improving points and deteriorating points (≥ 3) repeatable at 3 consecutive examinations
A	Based on Advanced Glaucoma Intervention Study (AGIS) method
A1	AGIS method
A2	CIGTS method
P	Based on point-wise linear regression analysis (PLR)
P1	PLR with threshold values, using a criterion of $p \leq 0.001$ and a cut-off point of ≥ 4 locations
P2	PLR with threshold values, using a criterion of $p \leq 0.001$ and a cut-off point of ≥ 3 locations
P3	PLR with threshold values, using a criterion of $p \leq 0.001$ and a cut-off point of ≥ 2 locations
P4	PLR with threshold values, using a criterion of $p \leq 0.001$ and a cut-off point of ≥ 1 location
P5	Two-omitting PLR with pattern deviation values, using criteria of ≤ -1 dB per year and $p < 0.05$ and a cut-off point of ≥ 3 locations at 3 consecutive examinations
P6	Two-omitting PLR with total deviation values, using criteria of ≤ -1 dB per year and $p < 0.05$ and a cut-off point of ≥ 3 locations at 3 consecutive examinations
P7	PLR with total deviation values, using criteria of ≤ -1 dB per year and $p \leq 0.01$ and a cut-off point of ≥ 3 locations at 3 consecutive examinations, with consistent worsening through the end of follow-up
P8	PLR with PROGRESSOR software (threshold values), using criteria of ≤ -1 dB per year and $p \leq 0.05$ and a cut-off point of ≥ 2 nonedge locations
I	Based on linear regression analysis with indices
I1	Linear regression analysis of a glaucoma hemifield cluster, using a criterion of $p < 0.005$ and a cut-off point of ≥ 4 clusters
I2	Linear regression analysis of a glaucoma hemifield cluster, using a criterion of $p < 0.005$ and a cut-off point of ≥ 3 clusters
I3	Linear regression analysis of a glaucoma hemifield cluster, using a criterion of $p < 0.005$ and a cut-off point of ≥ 2 clusters
I4	Linear regression analysis of a glaucoma hemifield cluster, using a criterion of $p < 0.005$ and a cut-off point of ≥ 1 cluster
I5	Linear regression analysis of the MD, using a criterion of $p < 0.025$
I6	Linear regression analysis of the MD, using a criterion of $p < 0.05$
I7	Linear regression analysis of the CPSD, using a criterion of $p < 0.025$
I8	Linear regression analysis of the VFI, using a criterion of $p < 0.05$
T	Based on PLR and linear regression analysis with indices
T1	Progression according to threshold noiseless trend (TNT) program
T2	Suspected progression according to TNT program
C	Based on clinical judgement
C1	Likely progression based on the subjective assessment of Humphrey printouts by multiple observers without using a predefined criterion; observers should reach agreement and were blinded for other clinical information
C2	Likely progression based on the subjective assessment of visual fields by multiple observers using PROGRESSOR software without a predefined criterion; observers should reach agreement and were blinded for other clinical information
C3	Nonparametric progression analysis based on nonparametric ranking of MD values
C4	At least one of these scotoma criteria: deepening of ≥ 2 points by ≥ 10 dB in the same locations; expansion of ≥ 2 points by ≥ 10 dB; development of ≥ 2 adjacent points, not within or adjacent to a baseline scotoma, with $p < 1\%$ or worse at the pattern deviation plot or progression of 1 previously normal location in central 10 degrees by ≥ 10 dB. Confirmations at the same locations on 3 consecutive fields, based on pattern deviation values
C5	At least one of these scotoma criteria: deepening of ≥ 2 points by ≥ 10 dB in the same locations; expansion of ≥ 2 points by ≥ 10 dB; development of ≥ 2 adjacent points, not within or adjacent to a baseline scotoma, with $p < 1\%$ or worse at the pattern deviation plot or progression of 1 previously normal location in central 10 degrees by ≥ 10 dB. Confirmations at the same locations on 2 consecutive fields, based on pattern deviation values

Methods are classified into six groups according to their main characteristics. GPA = glaucoma progression analysis, AGIS = Advanced Glaucoma Intervention Study, CIGTS = Collaborative Initial Glaucoma Treatment Study, PLR = point-wise linear regression, dB = decibels, MD = mean deviation, CPSD = corrected pattern standard deviation, VFI = visual field index, TNT = threshold noiseless trend.

Glaucoma Treatment Study (CIGTS) method, which is derived from the AGIS method, use a scoring system to grade each visual field in the follow-up period. The AGIS score is based on the actual decibel deviations at the total deviation plot, while the CIGTS algorithm is based on the p-values obtained from the total deviation probability plot. Both scoring systems

range from 0 to 20, with 0 representing no field loss and 20 end-stage disease. Visual field series are considered to be progressive if the score has a minimal increase of four (with the AGIS method) or three points (with the CIGTS method) and is confirmed by two additional tests (Katz 1999). In methods that are classified as point-wise linear regression (PLR)

methods (group P from PLR), a linear regression analysis was performed in different individual locations at the visual field. Another group of methods was based on a linear regression analysis with visual field indices (group I from indices), like the MD and the visual field index (VFI). The VFI is calculated by the new software of the HFA. Each location on the

visual field contributes to the VFI, although it is more heavily weighted to central areas of the visual field (Bengtsson & Heijl 2008). A location that is not significantly ($p < 0.05$) depressed on the pattern deviation probability map is considered to have a 100% sensitivity. The VFI is expressed as one percentage, where 100% represents a normal visual field and 0% represents a perimetrically blind eye. The HFA software performs a linear regression analysis of the VFI against time (Bengtsson & Heijl 2008).

Another group of methods was made that combined a PLR and a linear regression analysis of the MD value. The methods in this group (group T from TNT) were variants of the threshold noiseless trend (TNT) program (Gonzalez De La Rosa et al. 2009). The TNT program filters perimetric results and takes into account dependency relations in the visual field. Moreover, it combines linear regression analyses of the MD, the cumulative defect curve, and different locations at the visual field. Suspected progression is seen for the first time that one of these parameters indicates progression. If this result is repeated by two consecutive examinations or if two or more parameters indicate progression, TNT indicates definite progression (Diaz Aleman et al. 2009). Finally, a clinical group (group C from clinical) comprised methods based on clinical judgement. Classified in this group were methods based on entirely subjective assessments of visual fields by multiple observers, who had to agree on progression while they were blinded for other clinical data. Other methods in this group used certain algorithms for the assessment of visual fields, for example based on the clinical judgement of scotoma's (Boden et al. 2004). One clinical method was based on nonparametric ranking of MD values (Wesselink et al. 2009). This method objectifies the commonly practised method of monitoring glaucoma patients with the use of MD values. A visual field series is considered progressive as the MD value of a follow-up visual field is worse than the MD of the worse of two baseline fields. This has to be confirmed on at least two visual fields (Wesselink et al. 2009). We did not take into account

one subjective method that was based on the experience and expertise of only one observer who was not blinded for other clinical information (Birch et al. 1995).

Population characteristics

In total, 1040 eyes of 948 patients with glaucoma were studied in the ten studies included. As shown in Table 2, all patients were derived from western countries, with mean baseline MD values ranging from -3.3 to -10.4 decibel (dB), and mean age ranging from 57.7 to 72.5 years. An average of 1.7 visual fields per year were analysed in the studies. In one study, patients received laser therapy or filtration surgery according to a treatment protocol (Nouri Mahdavi et al. 2007). All other patients received various treatments during the follow-up period. An average of six methods was directly compared in the studies included. The uncorrected incidence proportions that were given or calculated for the different methods in the studies are also shown in Table 2. The maximum difference in the incidence proportions within a single study was 0.44.

Ranking of methods

Figure 2 shows the ranking of the 30 methods according to their incidence proportions when combined in the meta-analysis. The mean estimated incidence proportion of progression was 0.21 (95% confidence interval (CI) 0.15, 0.26) in 6 years (dashed vertical line in Fig. 2), indicating that on average 21% of the study eyes progressed in 6 years. The incidence proportions of progression according to 30 methods ranged from 0.02 (CI -0.02 , 0.05) for method P1 to 0.62 (CI 0.47, 0.78) for method P8. A continuous vertical line represents the incidence proportion of the GPA (method G1), which was the most frequently studied method, with six studies in this meta-analysis. With an incidence proportion of 0.16 (CI 0.14, 0.19), the GPA is in the middle of the ranking of all 30 methods. The AGIS-based methods and most of the methods based on linear regression analysis with indices showed lower incidence proportions than the GPA method. Methods that are based on clinical

judgement or the TNT program showed higher incidences than the GPA. All incidence proportions shown in Fig. 2 are the estimated rates when the mean follow-up time was set at 6 years and the mean baseline MD value was set at -7 dB. These values were chosen because these were the mean values in the ten included studies. Follow-up time was significantly associated with the incidence of progression, with an increase in the incidence proportion of approximately 2.1 per cent points per extra follow-up year ($p < 0.001$). Baseline MD was also significantly associated with the incidence of progression, increasing the incidence proportion by 0.9 per cent points per extra dB of MD loss ($p = 0.025$). No accelerating (quadratic) relationships between these two predictors and the incidence proportions were found ($p = 0.93$ and $p = 0.77$, respectively).

Contribution of methods

The variance explained by the model can be expressed in terms of the amount of heterogeneity as estimated by the model (i.e. the estimate of τ^2). After correcting for follow-up time and baseline MD, the methods used for assessing progression accounted for all the remaining heterogeneity between the estimated incidences of progression in studies (CI 0, 0.002). To estimate the contribution of the chosen method in the quantification of the incidence of progression, the model in which all methods were assumed to have the same average incidence of progression (without including mean follow-up time and MD value as covariates) was compared with the model where all methods were allowed to have different incidences. The former model yielded an estimated amount of heterogeneity (i.e. the estimate of τ^2) equal to 0.011 (CI 0.006, 0.022), which is decreased to 0.002 (CI <0.001 , 0.008) when all methods were allowed to have different incidences. Therefore, approximately 82% of the heterogeneity in this analysis can be accounted for by the variety of methods used in studies. The rest of the heterogeneity was explained by the mean baseline MD value and the mean follow-up time.

Table 2. Characteristics of the ten studies included.

Study (year of publication)	Source of the study population	Method	Uncorrected incidence proportion of progression	Maximum difference in incidence of progression	Number of patients and eyes	Mean baseline MD value (dB)	Mean age (years)	Mean follow-up period (years)	Mean number of visual field examinations	Number of visual fields per year
Birch et al. (1995)	Royal Liverpool University Hospital, UK	P8 I6	0.56 0.14	0.42	NA 36	-7.5	NA	3.3*	6	1.8
Smith et al. (1996)	Glaucoma Service, Wilmer Institute, Baltimore, Maryland, USA	P1 P2 P3 P4 I1 I2 I3 I4 I5 I7	0.05 0.07 0.11 0.19 0.06 0.07 0.09 0.18 0.49 0.14	0.44	191 191	-8.7	61.8	7.1	9.5	1.3
Katz (2000), Katz et al. (1997, 1999)	Glaucoma Screening Study, Baltimore, Maryland, USA	G1 G2 A1 A2 P2 P3 P4 I1 I2 I3 I4 I5 I7 C1	0.21 0.33 0.10 0.19 0.03 0.10 0.22 0.06 0.10 0.12 0.25 0.18 0.07 0.21	0.30	56 67	-7.4	62.0	6.3**	7**	1.1
Lee et al. (2002)	DIGS, University of California, San Diego (and other locations), USA	G1 G2 G4 G5 A1 A2	0.08 0.13 0.08 0.13 0.04 0.04	0.09	48 48	-4.7	57.7	2.8	5	1.8
Viswanathan et al. (2003)	Moorfields Eye Hospital, London, UK	C1 C2	0.19 0.48	0.22	27 27	-7.7**	61**	5.7**	16	2.8
Artes et al. (2005)	QEII Health Sciences Centre, Halifax, Nova Scotia, Canada	G1 G2 P5 P6	0.21 0.30 0.24 0.36	0.15	101 168	-3.3**	61**	10**	21**	2.1
Nouri Mahdavi et al. (2007)	Advanced Glaucoma Intervention Study, USA	G2 A1 P7	0.31 0.22 0.35	0.13	156 156	-10.4	64.4	8	18**	2.3
Arnalich Montiel et al. (2009) [#]	Ramon y Cajal Hospital, Madrid, Spain	G1 G3 I6 I8 C4 C5	0.19 0.31 0.21 0.13 0.19 0.30	0.18	90 90	-6.3	72.5	5.1	6.2	1.2
Diaz Aleman et al. (2009)	Aberdeen Royal Infirmary, Aberdeen, UK	G1 G3 T1 T2 C1	0.11 0.25 0.43 0.52 0.14	0.41	42 56	-3.9	68.5	5.5	7.8	1.4
Wesselink et al. (2009)	GLGS, UMCG, Groningen, the Netherlands	G1 C3	0.16 0.33	0.17	201 201	-9.4 ^{###}	66.4 ^{###}	5.3 ^{###}	7.1 ^{###}	1.3

All data were directly derived from the articles, but we calculated the number of visual fields per year by dividing the number of visual field examinations by the follow-up period. * = the follow-up period for the total group of patients with glaucoma and patients with ocular hypertension was used, ** = median value, # = data were derived from a more recent article (Casas Llera et al. 2009), which described the same study, ### = data for the total group of 221 patients was used. MD = mean deviation, dB = decibels, NA = data not available, DIGS = Diagnostic Innovations in Glaucoma Study, GLGS = Groningen Longitudinal Glaucoma Study, UMCG = University Medical Center Groningen.

Additional analysis

We evaluated the network model with both direct and indirect comparisons of methods by performing an additional analysis for the direct comparisons only. This analysis was also corrected for the mean follow-up time and the mean baseline MD. Methods G1 and G2 were the only methods that were directly compared in three or more studies. The estimated difference in the incidence proportion between methods G1 and G2 was 0.081 (CI 0.043, 0.120) in the model with only direct comparisons, while the difference in incidence was 0.077 (CI 0.033, 0.121) in the combined direct and indirect model. The direction and size of the difference between methods G1 and G2 was approximately the same in the direct model as in the combined direct and indirect model, although the combined model had a broader confidence interval.

Discussion

The mean incidence of glaucomatous visual field progression that we found in the literature was 21% in 6 years. This incidence was based on a large number of patients and a large number of methods to assess progression and can be seen as an average proportion of eyes that were considered to be progressive. The incidence of visual field progression had a large range from 2% to 62% in 6 years, depending on the method that was used. This again confirms the importance of the choice of a method to assess glaucomatous visual field progression. We then ranked these methods according to their incidence of progression. It is important to note that the incidences that we have found for each method can only be generalized for clinically treated glaucoma patients with a mean baseline MD value around -7 dB and a mean follow-up time of 6 years. The estimates of progression should be corrected by adding 0.9% to the incidence or by subtracting 0.9% from the incidence, for each dB decrease and increase in baseline MD value, respectively. In the same way, the incidences should be corrected by adding 2.1% for each year extra follow-up. The expected incidence of progression in an article that used another method can thus be estimated

with the ranking presented in this article.

Together with the range of incidences of progression, we also investigated to what extent the use of different methods will lead to differences in results between studies. The chosen method accounted for nearly all differences in the incidence of progression that we found in the included studies, with the exception of the part that can be explained by the baseline

MD value and the follow-up time. This underscores the importance of taking account of the method used when critically appraising and comparing studies on glaucomatous visual field progression. However, the present study also underlines the influence of baseline visual field loss and follow-up time on the incidence of glaucomatous visual field progression. When clinicians use the same method, these factors are even more important

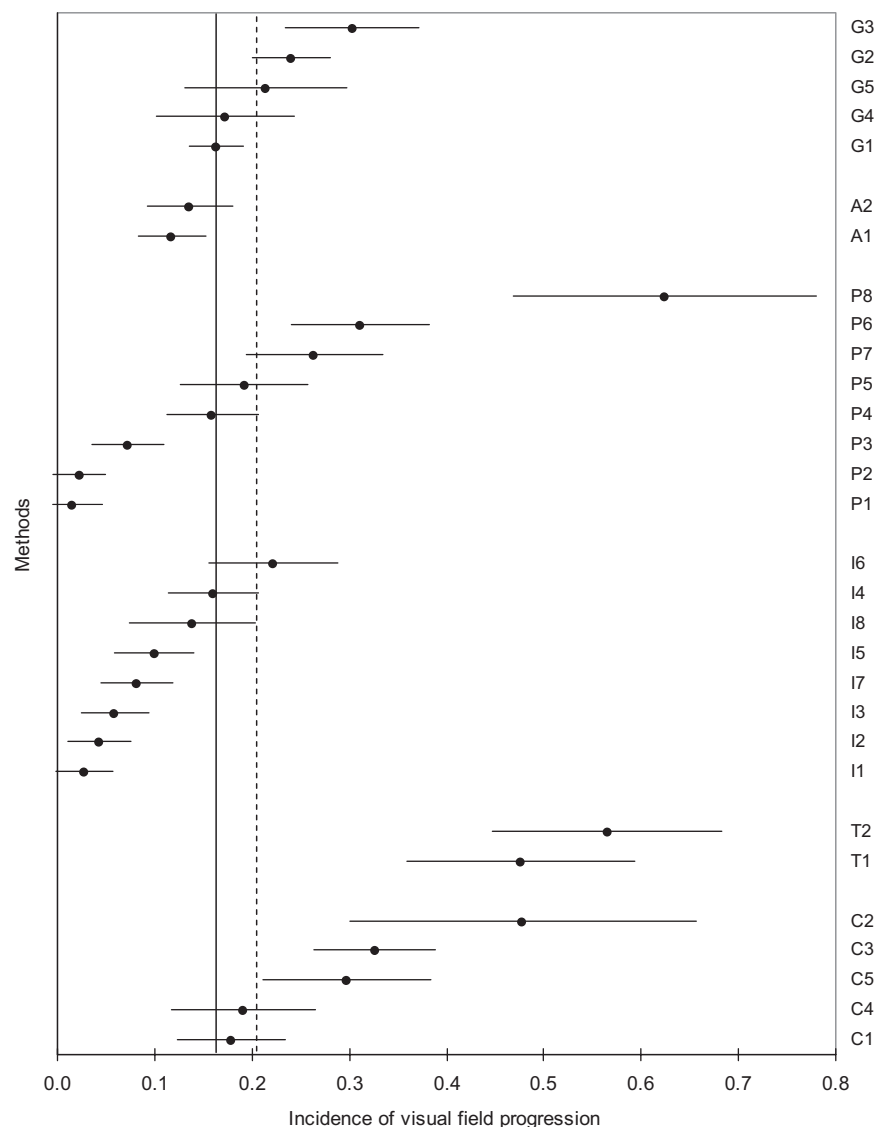


Fig. 2. A ranking of methods is shown according to their incidence of visual field progression. The 30 points represent the estimated incidences of progression with their 95% confidence intervals of all methods studied. The incidences are corrected for a mean follow-up period of 6 years and a baseline mean deviation value of -7 decibels. The dashed vertical line represents the mean estimated incidence of visual field progression in the ten studies. The continuous vertical line represents the incidence of the glaucoma progression analysis (method G1). The methods are grouped according to their main characteristics: group G = methods based on glaucoma progression analysis, group A = methods based on Advanced Glaucoma Intervention Study method, group P = methods based on point-wise linear regression analysis, group I = methods based on linear regression analysis with indices, group T = methods based on point-wise linear regression analysis and linear regression analysis with indices, group C = methods based on clinical judgement.

in clinical practice than the method that is used. Although a longer follow-up period logically leads to an increase in the incidence of visual field progression, a worse baseline visual field increases the chance of future visual field progression in an individual patient. It has been found in previous studies that more advanced baseline visual field loss was an independent predictor of future visual field progression (Leske et al. 2007). However, other studies have found a non-linear relation of baseline visual field and subsequent progression (Lichter et al. 2001; Polansky et al. 2003) or even an opposite relation (The AGIS Investigators 2000). The latter finding could be explained by floor effects because of advanced visual field loss of the AGIS patients at baseline. This means that no worsening can be observed when visual field loss is already advanced, but instead even an improvement of the visual field that is most likely the result of regression to the mean. With a mean baseline MD value of -7 dB, we assume that floor effects had no major influence on the present results.

Although most meta-analyses use only direct comparisons between treatments or diagnostic methods, we also used indirect comparisons. This 'network' meta-analysis is a method to quantify a rank order, even if no direct comparison between two methods has been conducted. Using this method, we made optimal use of all available data on visual field progression in patients with glaucoma. Differences in the incidence proportions that are found with direct comparisons can only be explained by the methods that are applied, because these comparisons are made in the same population. In contrast, indirect comparisons may also be influenced by population differences. Direct comparisons between methods can thus be seen as superior to indirect comparisons. However, in this meta-analysis, we corrected for follow-up time and baseline MD value, which reduced the unexplained heterogeneity between the incidences to zero. Therefore, we assume that the results of our meta-analysis provide an adequate approximation of direct comparisons between methods that are actually only indirectly compared. An additional analysis using only direct comparisons

resulted in approximately the same incidences as when both direct and indirect comparisons were used.

Remarkably, the incidences in the present study are much lower than the incidences of some other well-known studies. These studies were not included in our meta-analysis because they did not directly compare two or more methods. For example, 41% of the treated patients in EMGT progressed on their visual field in a follow-up period of approximately 6 years and with a baseline MD value of -4 dB (Heijl et al. 2002). The incidence that we found for the same method (i.e. the GPA method) was only 16%. This large difference could partly be explained by the frequent testing protocol of the EMGT, in which visual field tests were carried out at 3-month intervals (Leske et al. 1999). Because of the two required confirmations of the GPA method, the availability of at least four visual field tests per year leads to an increased chance to detect progression during the follow-up period. Moreover, a flexible testing protocol resulted in even more available visual fields in the EMGT, because additional visual field tests were performed within 1 month when tentative progression occurred. As shown in Table 2, this testing frequency is much higher than the mean frequency of 1.7 visual fields per year in the studies that were included in our meta-analysis. In addition, all studies in our meta-analysis only used a fixed testing protocol or no testing protocol. Furthermore, the large difference could also be explained by 61 patients in the EMGT who had two eligible eyes for the study. These patients had a higher chance to be progressive because they were considered to have progression when the first eligible eye met the progression criteria (Heijl et al. 2002). An additional explanation for the large difference could be that EMGT patients were not treated according to current standards. For example, no target pressure was defined for individual patients (Leske et al. 1999). A higher incidence of progression was also found in the Canadian Glaucoma Study (CGS) (Chauhan et al. 2008). The incidence of visual field progression in this study was 31% in 5 years in glaucoma patients with a median baseline MD value of -4 dB. In the

CGS, a variant of method G2 was used with a slightly stricter cut-off point. The estimated incidence of progression of method G2 was only 24% in our meta-analysis. In view of the better baseline visual field and shorter follow-up period in the CGS compared with the studies in our meta-analysis, one would just expect a lower incidence of progression in the CGS. However, the CGS also used a flexible testing protocol with a testing frequency of at least three visual fields per year and a required confirmation of suspected progression in 10 days (Chauhan et al. 2008). Overall, the studies in our meta-analysis better reflect current clinical practice than the above-mentioned studies. We found incidences that are based on a heterogeneous group of 948 patients. The mean frequency of visual field examinations of around 1.7 visual fields per year in the present meta-analysis also reflects a feasible testing frequency in clinical practice.

Our study gives insight into the methods that give either high or low estimates of the incidence of glaucomatous visual field progression but does not solve the problem of the lack of a gold standard. Therefore, the definition of 'true progression' remains a problem. A method may well result in a high incidence proportion without really differentiating between 'true progressive' and 'true nonprogressive' patients. A method that correctly classifies a large proportion of the 'true progressive' patients as progressive is said to be a method with a high sensitivity. On the other hand, when a method classifies a large proportion of the 'true nonprogressive' patients as nonprogressive, it is said to be a specific method. The choice of a cut-off point in a particular method is generally a trade-off between sensitivity and specificity. In general, we cannot say which method is the better one, but we know that sensitivities and specificities differ per method in a predictable way when another cut-off point is used. When the cut-off value in these methods is less strict, the sensitivity increases while at the same time, the specificity decreases. Examples in the present meta-analysis are the PLR-based methods P1, P2, P3, and P4, which only differ in the number of locations that should meet the same criterion. The criterion for

a location to be progressive is a statistically significant negative linear regression line ($p \leq 0.001$). Even without looking at the results of the meta-analysis, we can assume that method P1, which needs the highest number of affected locations for a visual field series to be progressive, is the least sensitive and the most specific. The same holds for methods I1, I2, I3, and I4, which differ only in the number of progressive clusters that are required to meet the definition of progression. It is important to notice that in epidemiologic studies, methods with a high specificity are preferred when the true incidence of progression is low. This leads to an estimate of the relative risk of the presumed prognostic factor that is closer to the true value of the relative risk. In clinical practice, a highly specific method with few false positives is also preferred when the treatment of progressive patients has considerable side-effects.

We used the results of 30 methods in our meta-analysis, for which a patient could only be progressive or nonprogressive. However, some methods were originally intended to quantify the rate of progression instead of a binary presence or absence of progression. These methods are based on indices of the complete tested visual field such as the MD (i.e. in methods I5 and I6) and the VFI (i.e. in method I8). In clinical practice, the MD index and the VFI are used to quantify the rate of progression on a continuous scale, ranging from severe disease progression (e.g. 16% loss in the VFI per year) to slight improvement (e.g. 3% increase in the VFI per year) (Bengtsson et al. 2009). Sole statistical significance of the slopes of these indices may be of limited interest for clinical practice, while knowing the rate of progression would direct appropriately aggressive treatment to those who are at highest risk for visual disability (Caprioli 2008). However, when a rate of progression is used to decide on changing therapy, it has to be decided which cut-off value for the rate leads to the decision to change therapy. Ultimately, even a rate method leads to a dichotomization of the population. Rates of progression should therefore be incorporated in progression criteria of methods to determine the presence or absence of progression. For example, progression

could be defined as a statistically significant decline of the VFI ($p < 0.05$) with a rate of at least 3% loss per year. Several PLR methods in the present meta-analysis (i.e. P5, P6, P7, and P8) already used statistically and clinically relevant criteria for individual visual field locations. Finally, the pattern standard deviation (PSD) or the corrected pattern standard deviation (CPSD) indices of the complete tested visual field (e.g. in method I7) should generally not be used in methods to assess glaucomatous visual field progression. Their values shift to normal values again in more advanced glaucoma (Heijl et al. 1987).

In conclusion, the incidence of glaucomatous visual field progression varies considerably between different studies. This is mainly caused by the variety of methods used to assess progression but also by differences in follow-up time and baseline visual field loss. Other factors, like the frequency of visual field examinations and the need for confirmation, also seem to be of importance. With a ranking of methods according to their incidence of progression, the interpretation and the possibility to compare studies are improved. Moreover, in several cases, the same method can be used with a more or less specific cut-off value, depending on the scientific or clinical goal. An ideal method combines statistically and clinically relevant criteria with tailor-made cut-off values.

Acknowledgements

No authors have any financial/conflicting interests to disclose.

References

Arnalich Montiel F, Casas Llera P, Munoz Negrete FJ & Rebolledo G (2009): Performance of glaucoma progression analysis software in a glaucoma population. *Graefes Arch Clin Exp Ophthalmol* **247**: 391–397.

Artes PH, Nicolela MT, LeBlanc RP & Chauhan BC (2005): Visual field progression in glaucoma: total versus pattern deviation analyses. *Invest Ophthalmol Vis Sci* **46**: 4600–4606.

Bengtsson B & Heijl A (2008): A visual field index for calculation of glaucoma rate of progression. *Am J Ophthalmol* **145**: 343–353.

Bengtsson B, Patella VM & Heijl A (2009): Prediction of glaucomatous visual field loss

by extrapolation of linear trends. *Arch Ophthalmol* **127**: 1610–1615.

Birch MK, Wishart PK & O'Donnell NP (1995): Determining progressive visual field loss in serial Humphrey visual fields. *Ophthalmology* **102**: 1227–1234; discussion 1234–1225.

Boden C, Blumenthal EZ, Pascual J, McEwan G, Weinreb RN, Medeiros F & Sample PA (2004): Patterns of glaucomatous visual field progression identified by three progression criteria. *Am J Ophthalmol* **138**: 1029–1036.

Bucher HC, Guyatt GH, Griffith LE & Walter SD (1997): The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* **50**: 683–691.

Caldwell DM, Ades AE & Higgins JP (2005): Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* **331**: 897–900.

Caprioli J (2008): The importance of rates in glaucoma. *Am J Ophthalmol* **145**: 191–192.

Casas Llera P, Rebolledo G, Munoz Negrete FJ, Arnalich Montiel F, Perez Lopez M & Fernandez Buenaga R (2009): Visual field index rate and event-based glaucoma progression analysis: comparison in a glaucoma population. *Br J Ophthalmol* **93**: 1576–1579.

Chauhan BC, Mikelberg FS, Balaszi AG, LeBlanc RP, Lesk MR & Trope GE (2008): Canadian Glaucoma Study: 2. risk factors for the progression of open-angle glaucoma. *Arch Ophthalmol* **126**: 1030–1036.

Diaz Aleman VT, Anton A, de la Rosa MG, Johnson ZK, McLeod S & Azuara Blanco A (2009): Detection of visual-field deterioration by Glaucoma Progression Analysis and Threshold Noiseless Trend programs. *Br J Ophthalmol* **93**: 322–328.

Gonzalez De La Rosa M, Gonzalez Hernandez M & Diaz Aleman T (2009): Linear regression analysis of the cumulative defect curve by sectors and other criteria of glaucomatous visual field progression. *Eur J Ophthalmol* **19**: 416–424.

Hardy RJ & Thompson SG (1996): A likelihood approach to meta-analysis with random effects. *Stat Med* **15**: 619–629.

Heijl A, Lindgren G & Olsson J (1987): A package for the statistical analysis of visual fields. In: *Doc Ophthalmol Proc Ser*. Dordrecht: Martinus Nijhoff Publishers 153–168.

Heijl A, Leske MC, Bengtsson B, Hyman L, Bengtsson B & Hussein M (2002): Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol* **120**: 1268–1279.

Heijl A, Bengtsson B, Chauhan BC, Lieberman MF, Cunliffe I, Hyman L & Leske MC (2008): A comparison of visual field progression criteria of 3 major glaucoma trials in early manifest glaucoma trial patients. *Ophthalmology* **115**: 1557–1565.

Katz J (1999): Scoring systems for measuring progression of visual field loss in clinical

- trials of glaucoma treatment. *Ophthalmology* **106**: 391–395.
- Katz J (2000): A comparison of the pattern- and total deviation-based Glaucoma Change Probability programs. *Invest Ophthalmol Vis Sci* **41**: 1012–1016.
- Katz J, Gilbert D, Quigley HA & Sommer A (1997): Estimating progression of visual field loss in glaucoma. *Ophthalmology* **104**: 1017–1025.
- Katz J, Congdon N & Friedman DS (1999): Methodological variations in estimating apparent progressive visual field loss in clinical trials of glaucoma treatment. *Arch Ophthalmol* **117**: 1137–1142.
- Lee AC, Sample PA, Blumenthal EZ, Berry C, Zangwill L & Weinreb RN (2002): Infrequent confirmation of visual field progression. *Ophthalmology* **109**: 1059–1065.
- Leske MC, Heijl A, Hyman L & Bengtsson B (1999): Early Manifest Glaucoma Trial: design and baseline data. *Ophthalmology* **106**: 2144–2153.
- Leske MC, Heijl A, Hyman L, Bengtsson B, Dong L & Yang Z (2007): Predictors of long-term progression in the early manifest glaucoma trial. *Ophthalmology* **114**: 1549–1572.
- Lichter PR, Musch DC, Gillespie BW, Guire KE, Janz NK, Wren PA & Mills M (2001): Interim clinical outcomes in the collaborative initial glaucoma treatment study comparing initial treatment randomized to medications or surgery. *Ophthalmology* **108**: 1943–1953.
- Mayama C, Araie M, Suzuki Y et al. (2004): Statistical evaluation of the diagnostic accuracy of methods used to determine the progression of visual field defects in glaucoma. *Ophthalmology* **111**: 2117–2125.
- Nouri Mahdavi K, Hoffman D, Ralli M & Caprioli J (2007): Comparison of methods to predict visual field progression in glaucoma. *Arch Ophthalmol* **125**: 1176–1181.
- Polansky JR, Juster RP & Spaeth GL (2003): Association of the myocilin mt.1 promoter variant with the worsening of glaucomatous disease over time. *Clin Genet* **64**: 18–27.
- Riley RD, Thompson JR & Abrams KR (2008): An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics* **9**: 172–186.
- Salanti G, Higgins JP, Ades AE & Ioannidis JP (2008): Evaluation of networks of randomized trials. *Stat Methods Med Res* **17**: 279–301.
- Schulzer M, Anderson DR & Drance SM (1991): Sensitivity and specificity of a diagnostic test determined by repeated observations in the absence of an external standard. *J Clin Epidemiol* **44**: 1167–1179.
- Smith SD, Katz J & Quigley HA (1996): Analysis of progressive change in automated visual fields in glaucoma. *Invest Ophthalmol Vis Sci* **37**: 1419–1428.
- Spry PGD & Johnson CA (2002): Identification of progressive glaucomatous visual field loss. *Surv Ophthalmol* **47**: 158–173.
- The AGIS Investigators (1994): Advanced Glaucoma Intervention Study. 2. Visual field test scoring and reliability. *Ophthalmology* **101**: 1445–1455.
- The AGIS Investigators (2000): The Advanced Glaucoma Intervention Study (AGIS): 7. The relationship between control of intraocular pressure and visual field deterioration. *Am J Ophthalmol* **130**: 429–440.
- Viswanathan AC, Crabb DP, McNaught AI, Westcott MC, Kamal D, Garway Heath DF, Fitzke FW & Hitchings RA (2003): Interobserver agreement on visual field progression in glaucoma: a comparison of methods. *Br J Ophthalmol* **87**: 726–730.
- Wesseling C, Heeg GP & Jansoni NM (2009): Glaucoma monitoring in a clinical setting: glaucoma progression analysis vs nonparametric progression analysis in the Groningen longitudinal glaucoma study. *Arch Ophthalmol* **127**: 270–274.

Received on March 15th, 2010.

Accepted on July 16th, 2010.

Correspondence:

Paul JG Ernest

University Eye Clinic Maastricht

PO Box 5800

6202 AZ Maastricht

The Netherlands

Tel: + 31 43 388 2242

Fax: + 31 43 387 5343

Email: p.j.g.ernest@gmail.com